NASA TECHNICAL TRANSLATION NASA TT F-15,593

SUBJECTIVE AND OBJECTIVE EVALUATION OF MACHINERY NOISE

M. Jahn

Translation of "Subjektive und objektive Bewertung von Maschinengeräuschen," Acustica, Vol. 16, No. 3, 1965-1966, pp. 175-186

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546 MAY 1974

STANDARD TITLE PAGE

1. Report No. NASA TT F-15,593	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle SUBJECTIVE AND OBJECTIVE EVALUATION OF MACHINERY NOISE		5. Report Date May 1974
		6. Performing Organization Code
7. Author(s) M. Jahn, Central Radio and TV Center of the German Post Office, Berlin-Aldershof		8. Performing Organization Report No.
		10. Work Unit No.
9. Performing Organization Name and Address Leo Kanner Associates, P.O. Box 5187, Redwood City, California 94063		11. Contract or Grant No. NASw-2481
		13. Type of Report and Period Covered Translation
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, WASHINGTON, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

Translation of "Subjektive und objektive Bewertung von Maschinengeräuschen," Acustica, Vol. 16, No. 3, 1965-1966, pp. 175-186.

16. Abstract

Subjective and objective evaluation of machinery noise; noises were judged subjectively by 28 investigators and were combined with several methods of objective estimation; from results, conclusions are drawn about accuracy and reproduction of subjective aural comparisons and about influence of results by measurements with different observer groups; subjective results were compared; testing of different calculation methods showed that for actual machinery noises, calculations of loudness from the third octave level according to S.S. Stevens leads to most accurate results.

17. Key Words (Selected by Author(s))	18. Distribution Statement Unclassified - Unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 29	22. Price 4.50

NASA-HQ

SUBJECTIVE AND OBJECTIVE EVALUATION OF MACHINERY NOISE

M. Jahn
Central Radio and TV Center of the German Post Office, Berlin-Aldershof

Summary /175*

Supplementary to an investigation by Lübcke, Mittag and Port on the subjective and objective calculation of machine noises, the same noise bands were fully investigated once again at the Berlin radio and television center (RFZ). The noises were judged subjectively by 28 investigators, and combined with several methods of objective estimation.

From the results, conclusions are drawn about the accuracy and reproduction of subjective hearing comparisons and the influence of the results of such measurements with different observer groups. The subjective measurement results of the institutions concerned were compared with one another under the assumption:

$$L_{1\ kHz} = L_{Normal} - 2\ dB.$$

The results from the RFZ lay about 2.3 dB ± 2.1 dB below the Stuttgart values and 0.3 dB ± 1.6 dB below those from TU Berlin.

The "testing" of the calculation methods in RFZ led to the result that for practical noises of a similar kind, the calculation of the loudness from the third-octave level according to Stevens leads to the most accurate results. For many practical cases in conisdering the diffusion range of subjectively ascertained loudness values from s = ±4 dB to ±5 dB, the use of octave analysis according to Niese or Stevens is regarded as sufficiently accurate. These conclusions are supported by the results from the TU Berlin.

*Numbers in the margin indicate pagination in the foreign text.

1. Introduction and Purpose /176

Lübcke, Mittag and Port have reported on "Subjective and objective evaluation of machinery noise" [1].

At the Institute for Technical Acoustics of the Technical University of Berlin, and at the Institute for Communications Technology of the Stuttgart Technical College, a figure for noise from internal combustion engines is judged subjectively under the most constant possible conditions, and combined with several methods of objective evaluation.

This should establish the accuracy of subjective measurements and prove the feasibility of objective experiments.

The Radio and Television Center of the German Post Office has conducted comparable experiments.

In this report we shall describe our experiments, discuss the results and compare them with the results reported by Lübcke et al [1].

A copy of the Stuttgart noise manuals was available for the RFZ measurements. Ten of the listed noises were selected.

In our experiments special importance was placed on proving the reproducibility and accuracy of such subjective measurements. The subjective comparisons are then used to evaluate the objective methods.

Port and Lübcke have reached agreement over measurement methods and apparatus [1]. These involve the frequency range of the communications chain, the measurement level, the normal noise level and the measurement methods. Individual measurements were conducted in accordance with these guidelines.

2. Subjective Measurement of Sound Intensity

Measurement of sound volume means, precisely, that a 1 kHz tone will be found to be just as loud as the volume of a test noise. In practice, one uses a narrow bandwidth of noise centered about 1000 Hz, instead of a pure sine wave. This renders the characteristics of the two sounds similar and so facilitates comparison. In the following investigations, the normal sound source was an AM narrow band noise centered at 1 kHz with a bandwidth Δf = 200 Hz and a peak amplitude of 4 $\underline{\Delta}$ 12 dB. This was produced from white noise by passage through a bank of filters with third-octave and octave filters to isolate the desired band. Figure 1 shows the frequency distribution of the normal noise, plotted from measurement points.

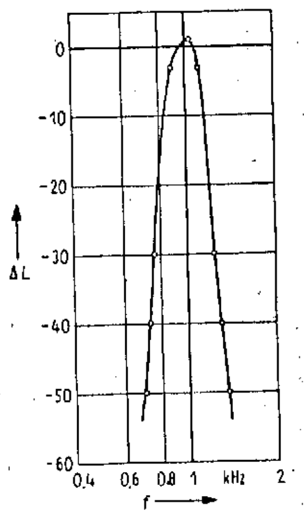

Fig. 1. Frequency distribution of the normal or standard noise at the measurement point.

According to the theory of frequency groups [2], the level of the standard used had to be slightly lower in order for the levels of the 1 kHz tone and the frequency ensemble centered at 1 kHz (Δf = 160 Hz) to exhibit the same loudness. A loudness comparison between the narrow bandwidth noise used and the 1 kHz tone showed that the sine wave at the same level was 2 dB louder (see also Sec. 2.7).

Niese [3] has demonstrated the same result independent of sound level in the comparison of the frequency ensemble noise with the 1 kHz tone. If these results are accepted as well-founded, this indicates that in all investigations to date in which an AM narrow band noise (Δf = 160 Hz) was used as a substitute of identical level for the 1 kHz standard tone, the subjective sound levels were reported at around 2 units too high. This result was considered in comparison with our findings (see Sec. 4).

2.1. Research Program

It should be demonstrated by specifically designed experiments what degree of precision is possible in principle in subjective loudness measurements, and what effect the selection of research subjects can have on the measurement results.

To this end, the individual diffusion ranges for the various judgments of the same noise, as well as the diffusion ranges of findings of various groups, were ascertained.

2.2. Measurement Principle /177

The method of pendular comparisons (method of tracking) should be employed for the measurements. Considering the apparatus available, the automatic level regulator must be set by a hand-controlled radio device, in accordance with production procedures. In other words, the research subjects adjust the variable level themselves from "seems too loud" or "seems too soft" toward the final value of "equally loud."

The standard sound intensity of the noise was measured against a level L = 74 dB, and the object sound intensity against a standard or normal level L = 80 dB (the relative sound intensities indicate which of the two sounds will be regulated).

As a final result the interpolated sound intensities will be evaluated as the average of the two.

2.3. Measurement Conditions

The echoless studio of the RFZ served as the measurement room. Listeners were situated approximately 1.5 meters in front of the loudspeakers. A rigidly placed chair with a headrest ensured an identical distance from the noise source for all measurements and prevented motion of the head away from the sound path.

The lowering of the level by 6 dB below the Stuttgart proposed value was necessary technically, in view of the low tolerance of the research subjects. With the measurement loudspeaker installed in the room, the requisite level of the normal tone routinely was not produced without distortion. For the normal tone level reached extreme values during the regulation process, up to 25 dB above the noise level.

A larger expenditure for new measurement equipment was not deemed necessary, as the aural characteristics of the level range in question do not change in practice.

The frequency range of the reproduction network was limited to a band between 100 Hz and 10 kHz by commercial high- and low-pass filters.

A copy of noise band spectra from the Institute of Communications Technology of the Stuttgart Technical College was used to select the test noises with level tone bursts and noise durations less than one minute at our disposal. For a ten-noise series, the most constant level possible was superimposed, and a measurement band with a noise duration of about 3 minutes was produced through a wider section. The band apparatus was measured along with the superimposed level tone bursts.

2.4. Research Subjects

The observer group consisted of 28 persons, some of whom were trained in subjective aural comparison and some of whom were untrained. A pure-tone audiogram was taken from all the participants before the start of the measurements, according to the Békésy audiometer principle, to weed out persons with hearing handicaps. To improve the audibility of the measurement tone, it was transmitted in impulses of about 200 msec.

The ages of the observers ranged from 25 to 37.

2.5. Measurement Procedure

The research subject was exposed via the loudspeaker to the normal and test noises alternating in the rhythm 1.5 sec noise -- 0.5 sec pause -- 1.5 sec noise. One of the two remained constant in intensity over the range; the other was controlled over the same loudspeaker. The observer was thereby able to perceive the sound that was clearly of lesser intensity, and to adjust it slowly toward the final value. Figure 2 shows an example of this control procedure. The participants of such experiments must always be given enough time to carry out the measurement procedure. Each measurement pattern must open with a particularly simple test -- the comparison of normal tone / normal tone. The same test serves to end a series, as a control of evaluation reliability. The test level employed can be varied from comparison to comparison by damping the regulator channel, to prevent a test subject from "getting used to" a control knob position and consequently no longer concentrating on achieving a valid measurement. The start notice proceeded through a rapid lowering of the level. A comparison lasted on the average 1 minute, and the recovery pause between two measurements, needed to prepare the next band and /178
readjust the standardization and damping factor, was about as long. To prevent fatigue it was necessary that a measurement session last no longer than 15 or 20 minutes. A measurement set thus consisted of no more than eight noise bursts. The entire level process was recorded from input to output, following painstaking standardization

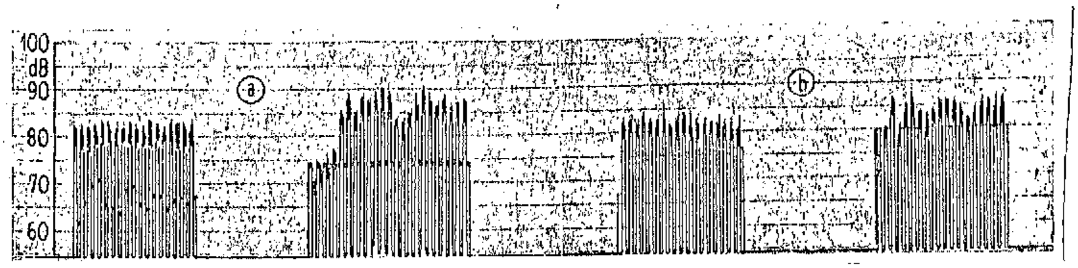

Fig. 2. Examples of the recording of the regulation procedure in the measurement of the standard and object sound intensities. Paper velocity 0.3 mm/sec, pen velocity 100 mm/sec. (a) Standard -- machinery noise; (b) standard -- 1 kHz.

of the apparatus with a Brüel & Kjaer Type 2305 level recorder.

2.6. Test Structure

Figure 3 shows the block diagram for the measurement of the standard and object sound intensities. To juxtapose machinery noise and the normal ensemble, two studio recorders are used. Before reaching the required control devices and amplifiers, the channels are fed through an electronic switching device. The following combination of high- and low-pass filters trims the signal range, and a high quality measurement loudspeaker completes the "transmission belt."

A Neumann Corp. MM2 served as a measurement microphone. The frequency range was verified by the comparison method through standardization against a Brüel & Kjaer measurement microphone. The measurement process was completed with the Brüel & Kjaer Type 2112 sound frequency spectrometer and the Type 2305 level recorder. Effective values can be measured with both devices. The overload range is 11.5 dB. The studio amplifier used has a small reserve. So when the noise peak amplitude factor climbs to 5, the total measurement series is driven at no more than 12 dB under full output, the analyzing equipment no more than 3 dB.

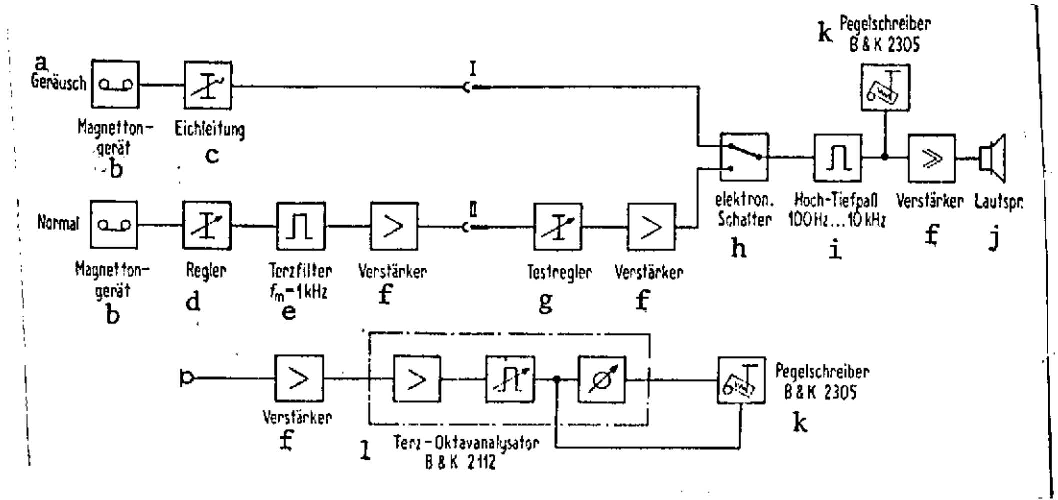

Fig. 3. Block diagram of the measurement apparatus.

Key: a. Noise
b. Magnetic sound source
c. Level control
d. Regulator
e. Third octave filter
f. Amplifier
g. Test regulator
h. Electronic switch
i. High/low pass
j. Loudspeaker
k. Level recorder
l. Third-octave/octave analyzer

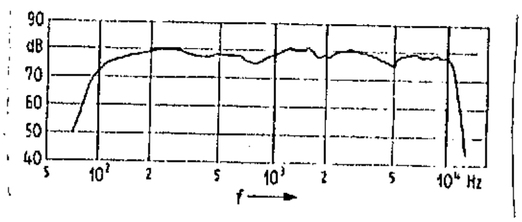

Fig. 4. Frequency distribution of the total measurement series.

The total frequency distribution of the measurement series from the output of the sound band equipment over the loudspeaker, the microphone, up to the level recorder is shown in Fig. 4. It was measured with automatic frequency feed with pure sinusoidal sound.

2.7. Results of the Subjective Measurements

2.7.1. Investigation of Accuracy Possibilities -- Individual and Group Variations

To determine the judgment reliability of individual research subjects and the group variations, a group of 15 observers next was required to judge only three noises. Each research subject was allowed ten judgments for each noise (spread over at least 5 different days). This experiment was limited to the measurement of the standard sound intensity levels.

Four intermediate values were recorded for each noise, namely after

each 1 judgment per research subject,
each 2 judgments per research subject and in toto,
72 judgments per 15 research subjects
and 100 judgments per 15 research subjects.

The total frequency percentages are calculated from these groups of values, and plotted on a probability graph with coordinates distorted according to the gaussian distribution, on which the measurement values appear as a straight line. The resulting curve of plotting points can be well approximated by a straight line. The standard deviation can thus be computed by gaussian distribution-weighted coefficients:

$$s = \pm \sqrt{\frac{\sum^{\nu} (\Delta L_{\nu} - \overline{\Delta L})^2}{n-1}}$$

Confidence range m = ±s/n.

These four curves for noise are shown in Fig. 5. This shows /179 that the distribution range, known from the slope, and independent of the number of measurement points for the selected observer group, remains approximately constant. The standard deviation for 72 measurements was ±3.5, ±4.8, and ±4.3 dB for the first three noises. The reliability range lies in the order of magnitude of ±0.5 dB.

For the comparison between individual distribution ranges and the group tolerances, finally the results of a group of 11 persons were evaluated. Ten judgments per noise were obtained from each observer.

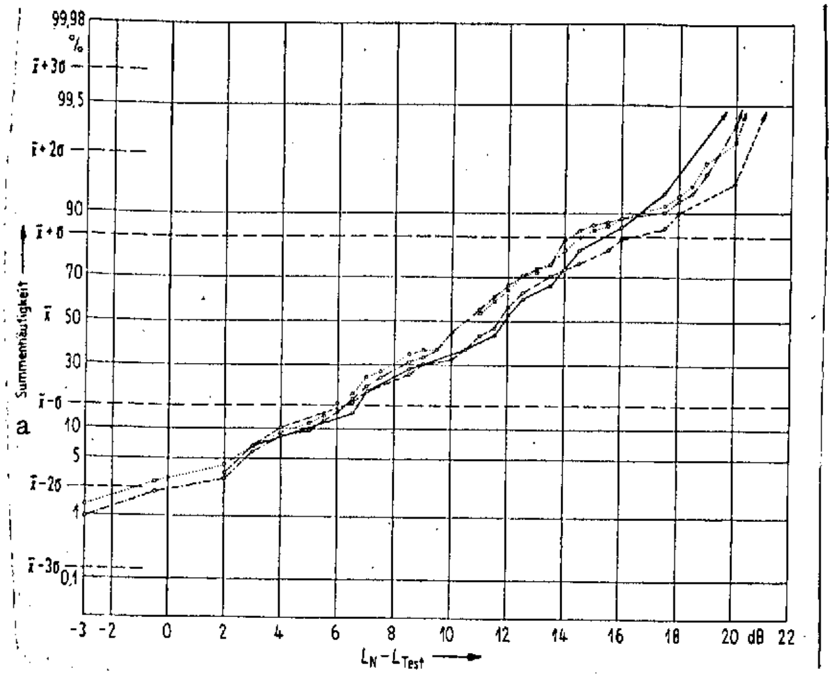

Fig. 5. Distribution frequency of the results of numerous judgments of the noise D 8 by 15 research subjects.

—— 1 judgment per research subject
— — 2 judgments per research subject
--- 72 judgments per 15 research subjects
-·- 100 judgments per 15 research subjects

Key: a. Cumulative frequency

Once there were large deviations because of extraordinary measurement conditions (severe cold, boredom, and the like). Ignoring these exceptional cases, a standard deviation of s = ±2 dB can be calculated as the average for the measurement values of individual observers.

This spread is very small. However, the mean values found by different observers lie up to 14 dB apart. This is not evidence of uncertainty in the research findings, but rather a fundamentally different loudness responsivity among different research subjects.

The group spread of the 11 persons lies for each sound in the order of magnitude of $s = \pm 4$ to ± 5 dB. It makes little difference in practice whether the test subject makes one judgment or ten. For the noise D_8, both values were calculated. The results were:

$$s_1 = \pm 4.56 \text{ dB}, \quad s_2 = \pm 4.4 \text{ dB}$$

The mean level difference between normal distribution and the test noise at the same intensity was rather small with sufficient practice. It reached a minimum of around 0.5 dB after ten measurements.

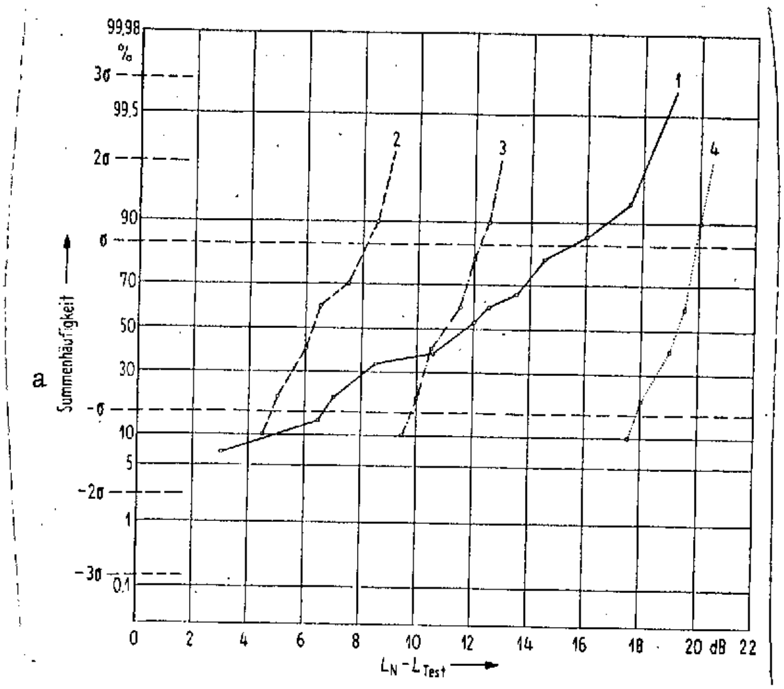

Fig. 6. Scalar frequencies of measurement values for the 11-member group with one judgment apiece (curve 1) and for three observers with ten judgments each (curves 2 to 4).

Key: a. Cumulative frequency

Figure 6 shows the frequency of the measurement values for the 11-member group with one judgment apiece, and for three observers with ten judgments each. The display shows clearly that the customary tolerance range for subjective tests is determined essentially through the tolerance range for the sound intensity findings, in which each observer comes up with good reproducible results. The search for an explanation for the different reactions of the observers to the wide band test noise based on absolute hearing thresholds, remains to be pursued. Test subjects with similar hearing thresholds produce widely scattered sound intensity measurements.

The sound intensity comparisons between the regulable and fixed narrow band noises which open and conclude each measurement series produce very closely fitting results. Here the individual conclusions converge:

$$\overline{\Delta L} = -0.1 \text{ dB}$$
$$s = \pm 0.7 \text{ dB}$$
$$m = \pm 0.08 \text{ dB}$$

(The specified level difference is a mean value of 97 separate measurements.)

2.7.2. Influence of Different Observer Groups on the Results /180

It will be clear from the foregoing that the discretion of the test subject plays a role in the results of such subjective measurements. The observer group must thus be representative for a large number of observers, if the results are to achieve general validity.

The group was then enlarged to 28 persons. Ten noises were now tested.

We sought to establish whether and to what extent the measurement results and tolerance width would be changed by enlarging the group.

For all sound intensity conditions, the frequency distribution was controlled by measurement of the standard sound intensity, specifically:

-- for the first 11 persons,
-- for all 28 persons, and on one trial:

- for a subgroup of seven persons, and
- for a subgroup of ten persons.

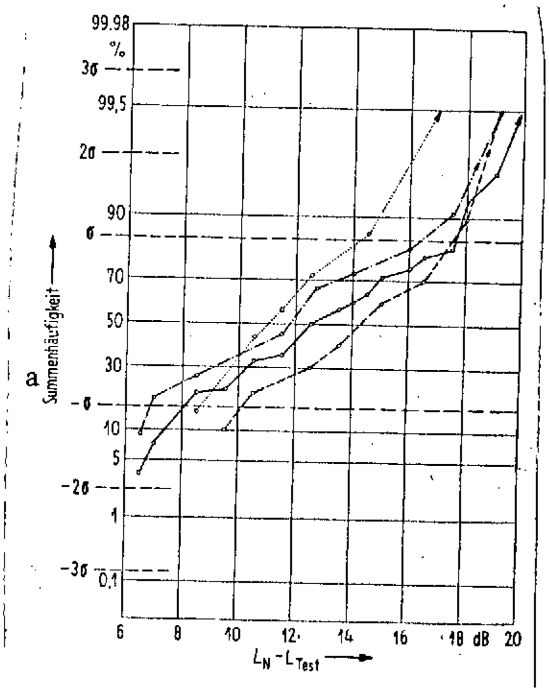

Fig. 7. Distribution frequency of the results of judgment of a noise by various observer groups.

-·- per 1 judgment for 11 research subjects
— — per 1 judgment for 10 research subjects
--- per 1 judgment for 7 research subjects
—— per 1 judgment for 28 research subjects

Key: a. Cumulative frequency

Figure 7 shows the frequency distribution of the measurement values for the four specified groups. From this one infers that with fewer test subjects, already a very acceptable distribution of measurement values is attained, and thereby for the groups concerned, a high degree of accuracy of the results is achieved. However, one sees also that different groups can arrive at results that differ by more dB. One concludes from these observations that it so happens that in such subjective hearing comparisons a small observer group of about ten persons can only be admissible, when specified pre-selection tests have determined that this group is an acceptable substitute for a larger number of persons, and when the mean value of the individual results is not essentially displaced by broadening the participating circle.

Increasing the group from 11 to 28 persons in no case /181
lowered the scatter s. It stayed constant or was about 1 dB larger.

The results of the first group of 11 persons were around 1 dB below the final values. This difference is insignificant, if one considers that the first group gained considerable practice during the course of the measurement program, so that the judgment accuracy was greater than that of the new additions [to the group]. This means that these 11 persons accordingly exhibited a good cross section of the optimum observer type. No increase in accuracy worth mentioning was expected from a subsequent enlarging of the group. For acceptability, the mean figures found with 28 persons showed the number of observers to be satisfactory. A satisfactory frequency distribution was obtained for all tested noises.

2.7.3. Final Results

The object sound intensity was supplementarily determined with 28 test persons for all ten noises, and the standard and object sound intensities of the normal employed.

All measured sound intensity segments are shown in Fig. 8 as differences in level between normal noise and test noise. The standard deviation is given. The mean difference between standard and object sound intensity varied within the observer group from 0.95 to 7.65 dB. The test subjects with the largest average values, 7.2 and 7.65 dB, were not excluded from the evaluation; for one of them, the standard sound intensity was always greater; for the other, the object sound intensity was always greater. Except for these two cases, the difference was always less than 5 dB, on the average 3.2 dB. The group average values of both sound intensities differed by 0.2 to 1.3 dB, the exceptions showing the noise K 10 with a deviation of aroud 4 dB.

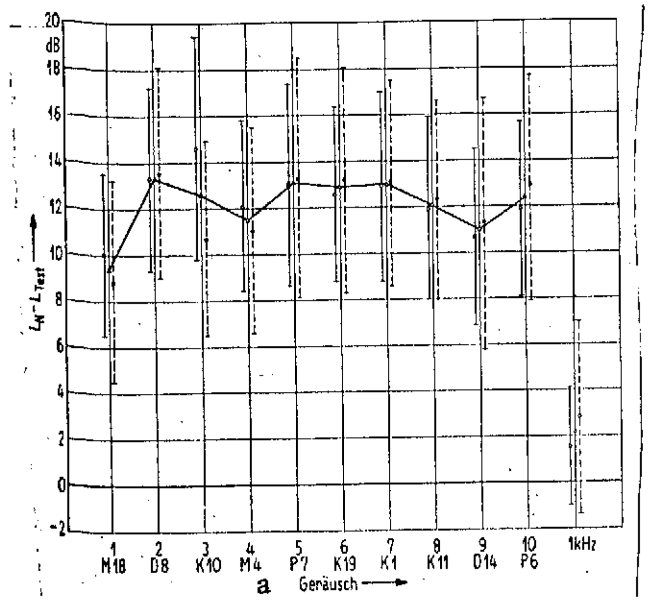

Fig. 8. Level differences between normal and machinery noises at the same sound intensities.

—— Normal sound intensity regulable (≙ standard sound intensity)
--- Normal sound intensity fixed (≙ object sound intensity)

Key: a. Noise

Mostly the differences in level between normal and machinery noise, and also the standard deviation with regulation of the test noises were larger:

$\bar{s}_{Stand.} = \pm 3.9$ dB

$\bar{s}_{Obj.} = \pm 4.6$ dB

With noises whose sound levels are somewhat higher, the maximum of the spectrum distribution is shifted toward higher frequencies.

A systematic influence of impulse persistence on the subjective sound intensity of the findings was not to be found.

In addition, Fig. 8 shows the findings of the normal noise/ /1 kHz comparison. The determined noise level difference of about 2 dB is a mean value of 50 measurements for 28 test subjects. For positing the sound intensities of machinery noise as the level for the equally loud 1 kHz tone, the established measurement values would be lowered correspondingly by 2 dB.

These results were also corrected by comparison with the results from [1].

3. Objective Evaluation of Noise

3.1. Measurement of the Effective Third-Octave or Octave Level

The objective measurement of the effective third-octave and octave level diagram shows the basis for the application of the sound intensity evaluation process. The corresponding measured evaluated levels, L_A and L_B, are commonly used for rough judgments of noise.

Measurements were made over the measuring microphone MM2 in the previously described observer chair at the measuring point in the headrest. Throught the use of the effective-value measuring-level recorder as recording instrument, reading the changing level values was considerably facilitated. The recording speed was 100 mm/sec. With this, the trace of a precision noise level measurement ($\tau \approx 94$ msec) was transcribed. The combination of the recorder with the spectrometer made possible an automatic indication of the spectrum and the unevaluated and evaluated total noise level. Each third octave was recorded in approximately 15 sec, each octave, in approximately 45 sec.

In Table I are displayed all octave levels and in Table II, /182
all third-octave levels of the noises for a total noise pressure level of 74 dB at the measurement point. To make possible a comparison with the third-octave levels measured in Stuttgart and Berlin, these spectra must be raised around 6 dB,and the corresponding deviations are listed in Table II as ΔL values. The effective third-octave levels are uniformly good, so far as is possible with the employment of different converters for noise reflection. Figure 9 shows the principal course of all spectra.

3.2. Evaluation of Noise

The following procedures are employed for the accurate setting of sound intensities and for rough evaluation of noises:

TABLE I. COMBINED CHART OF MEASURED EFFECTIVE OCTAVE LEVELS IN dB.*

Noise	Mean Frequency								
	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
M 18	< 43	63	71,5	68,5	63,5	61	55,5	46	
D 8	< 43	68,5	66,5	66,5	62,5	65,5	63,5	57,5	38,5
K 10	45,5	62,5	67,5	70	65,5	65	60,5	53	39
M 4	< 43	61	70,5	64,5	64	68,5	59	46,5	34,5
P 7	46	61	66,5	66,5	68,5	69	63,5	56	38,5
K 19	< 43	53	63	64	67	69	67	63,5	47
K 1	< 43	60	65	64,5	67	68,5	68	65,5	44,5
K 11	44,5	67,5	70	61	64,5	64,5	61,5	54,5	37,5
D 14	44	55,5	62	67,5	70	67	62	57,5	37,5
P 6	< 43	56,5	63,5	68,5	67,5	68	61	54,5	36,5

*[Note: Commas are equivalent to decimal points in all tables.]

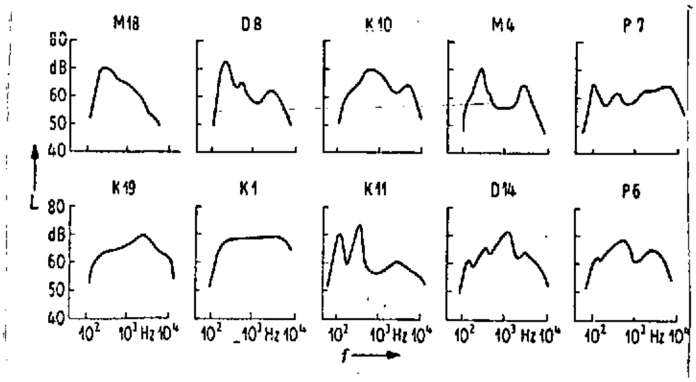

Fig. 9. Principal frequency distributions of all spectra.

a) Calculation of the sound intensity from the objectively measured third-octave spectrum following the proposal of Zwicker [4].

b) Calculation of the sound intensity from the objectively measured third-octave and octave levels according to the Mark-IV process of Stevens [5].

c) Calculation of the sound intensity from evaluated third-octave and octave levels following Niese [6].

d) Evaluation of noise with the effective noise pressure levels L_A and L_B. /183

e) Evaluation of the noises with NR-values.

TABLE II.

Noise		Mean Frequency in Hz from 1000 Hz to 10 kHz: 80	100	125	160	200	250	315	400	500	630	800	1	1.25	1.6	2	2,5	3.15	4	5	6,3	8	10	12,5	16	Noise Level L_A	L_m	L
M 18	a	35	48	60	61	68	67,5	63,5	65.5	63,5	61	57,5	58,5	59,5	56	56.5	57,5	54	48,5	45,5	44	40	37			69,5	73	
	b	+4	+2	0	−1	−0,5	+0,5	−1,5	−2 5	0	+1	+4	−2,5	−1.5	−2,5	−5,5	−5,5	−4	−3,5	−3,5	−3	−2	0			+0,5	0	74
	c	+5	+2	−1	−2	−3	−1,5	+0.5	+1,5	+1,5	+2	+5,5	−0,5	+0.5	+1	−2,5	−4,5	−2	+0,5	+0.5	0	+2	+5			0	0	
D 8	a	34	58.5	67,5	63	65	61	59.5	64	58,5	62,5	56,5	59	57	58,5	61	63.5	62,5	57,5	54,5	55	52	49,5			71	72,5	
	b	+4,5	−1,5	+2	+1	0	+0,5	+1,5	−2	+0,5	+0,5	+2,5	−2,5	−0,5	−2	−5	−4.5	−4,5	−4.5	−4.5	−4	−5	−4			−1	0	74
	c		−1,5	+1.5	−1	−2	−1	+3,5	+1	+1,5	+1,5	+2.5	+1	+1	−1,5	−3	−3,5	−3 5	−0,5	−1.5	−1	+1	+2.5			−1	0	
K 10	a	40	47	59,5	59,5	62	61,5	64,5	66.5	65.5	64,5	60,5	61,5	60	56,5	59,5	62,5	59,5	53,5	50,5	50.5	47,5	46.5			71,5	73,5	
	b	+11,5	+1,5	+2.5	+1	+0,5	0	−1	−2	+0,5	+0,5	+4	−1	−1,5	−0,5	−4	−4	−3,5	−2,5	−2	0	+5	+8			+0.5	0	74
	c	+7	+2	+0.5	−2,5	−3	−1,5	+0.5	+0,5	+0,5	+0.5	+3,5	+0,5	−1	+0,5	−2,5	−3,5	−4,5	−1,5	−0.5	−1,5	+8,5	+4,5			−0,5	−0,5	
M 4	a	34	50.5	54.5	58	56,5	70,5	60,5	59	60	59	56	60	59	57,5	58	66.5	55,5	51	53	44.5	40	34			72	73	
	b	+8	+3,5	+4,5	+3	+2.5	+0,5	0	0	+2	+2.5	+5	0	−1	−1 5	−2	−6,5	−2	−1	−3	−0,5	+1	−5			−1	0	74
	c	+17	+3,5	+2.5	0	−0,5	−0,5	+1,5	+2	+2	+3	+5	+1	+1	+0,5	+2	−2,5	−1,5	0	+1	−0,5	+2	−6			−2	0	
P 7	a	40	58	48	57	63	60.5	59	57,5	62,5	62	60.5	62,5	65	62,5	63	65,5	61,5	59	54	53,5	51	47			73	72,5	
	b	+2	+3	+3	+2	−2	+0,5	0	−0,5	+0,5	+3.5	+4.5	−0,5	−1	−1.5	−4	−4	−2,5	−3	−3	−1,5	+0,5	+3			0	+1,5	74
	c	+4	+2	+1	−1	−2	−1,5	+1	+0,5	+0.5	+2	+3,5	+1,5	0	+0,5	0	−3,5	−2,5	−1	0	−1,5	+2	+3			−1	+1,5	
K 19	a	35	42	48,5	61	56	59	59	59,5	59,5	60	60	63	64	63	63	66	64	62,5	61,5	60	59	57,5			74	73,5	
	b	+11.5	+4	+4	+2	+2	+1,5	+1	+0,5	+2.5	+3	+4	−1	−2	−2	−3	−3,5	−2,5	−3 5	−3	−1	0	+3,5			−0,5	+0,5	74
	c	+9.5	+3	+1,5	−1	−1	0	+2	+3,5	+4.5	+3	+4	+1	0	−1	−1	−3	−3	−1,5	−0,5	−1	−1	+1,5			−1	+0,5	
K 1	a	35	46	52	57,5	60,5	59	58	58	60,5	58	60	62	62	61,5	63	64	62,5	61	63,5	60	58	61			73	73	
	b	+9	+4	+4	+2.5	+2,5	+2	+1	0	+2,5	+4	+3	−0,5	−1	−1	−3	−3	−2,5	−2	−3,5	−1,5	−0,5	+2			−0,5	−0,5	74
	c	+7	+2	+2	−0,5	−0,5	0	+2	+2	+2,5	+3	+3	+2	+1	−0.5	−1	−3	−3,5	−1	−0,5	−2	0	+2			−1	−1	
K 11	a	34,5	57.5	66.5	61,5	58.5	69.5	62	55	56	58	54	57,5	63	58,5	59	61,5	59.5	57,5	52	51,5	51,5	49			70	72,5	
	b	+5.5	+0,5	+1,5	+1	+1	0	−1,5	−2	0	+1	+3	−1.5	−1	−3	−4,5	−5	−4.5	−4,5	−4	−2,5	−0,5	+4			−0.5	0	74
	c	+5,5	+1,5	+1	−0,5	−0.5	−0,5	0	+2	+1	+3	+4	+1,5	−1	+0,5	−1	−3.5	−2,5	−1	+1	−1.5	+0,5	+4			0	+0.5	
D 14	a	29,5	44,5	50,5	54	51.5	57	59,5	64,5	61	63	65	66,5	65,5	61,5	62	63,5	59	56.5	56	54,5	52	46,5			73	73.5	
	b	+12,5	+2.5	+4	0	+1,5	0	−0.5	−1,5	+1	+2,5	+4	−1,5	−2	−3	−5	−5	−5	−6	−8	−1.5	−10	−9,5			0	+0,5	74
	c		+1,5	+2.5	−3	−2,5	−3	−0,5	+1,5	0	+2	+3,5	+0.5	−1,5	−1,5	−3	−4,5	−5	−3.5	−3	+1,5	−4	−1,5			−0,5	+0,5	
P 6	a	31	41	51.5	54.5	56,5	56	60	62.5	63	64	60	62.5	63,5	62.5	63,5	63	58	55	52	52	49	43			72,5	74	
	6	<7	+3	+2.5	+3	+1,5	+1	−1	−0.5	+2.5	+3	+5.5	−0.5	−0,5	0	−2,5	−3	−0,5	−2,5	−2	−2	+1	+4,5			+0,5	0	74
	c		0	−2,5	+0,5	−4,5	−2	−2	+2,5	+1	0	+5	+0,5	+0.5	+0,5	−0,5	−2	0	−1	−2	−2	−1	+1			+1	0	

a. Effective third-octave level in dB.
b. Deviations ΔL from Port's values.
c. Deviations ΔL from Lübcke's values.

The results of all processes are displayed in Fig. 10 and Table III.

TABLE III. COMBINED CHART OF THE RESULTS OF ALL SOUND INTENSITY EVALUATION PROCEDURES CONDUCTED.

Quantity	Unit		Noise M 18	D 8	K 10	M 4	P 7	K 19	K 1	K 11	D 14	P 6	Normal
L	dB		74	74	74	74	74	74	74	74	74	74	80
L_N	phon		81,5	85,5	84,5	83,5	85	85	85	84	83	84,5	77,8
		a	+ 2,5	− 2,5	+ 3	+ 1,5	+ 3,5	+ 4	+ 5	+ 3	+ 0,5	+ 2	
		b	+ 3	− 3	+ 0,5	+ 0,5	+ 1	+ 1	+ 0,5	+ 1	− 0,5	− 1	
L_{NGF}	phon		87,5	90,5	89,5	88,5	90,5	91,5	91	89,5	90	89	
		a	− 1,5	− 2,5	− 1	− 1	− 2	− 1	− 0,5	− 2	− 2,5	0	
		b	− 1	− 1,5	− 1,5	− 1	− 1,5	− 1,5	− 1	− 1	− 2	− 1	
L_{NOD}	phon		80	83	82	82,5	84	84,5	85,5	82	83	83	
		a	− 1	− 2	− 0,5	− 2,5	− 1,5	+ 1	− 0,5	− 2,5	− 1,5	− 1	
		b	+ 0,5	0	0	0	0	+ 0,5	0	− 0,5	− 0,5	− 0,5	
L_{NTD}	phon		81	84,5	84	83,5	85	86	86	83	84,5	83,5	
L_{NIO}	phon		82,5	84	86,5	84	83,5	85,5	84,5	82	83,5	87,5	
L_{NIT}	phon		82	83,5	85	82	85	84,5	84	81,5	82	89	
L_A	dB		69,5	70,5	71,5	72	73	74	73	70	73	72,5	
		a	+ 0,5	− 0,5	+ 0,5	− 1	0	− 0,5	− 0,5	− 0,5	0	− 0,5	
		b	0	− 0,5	− 0,5	− 2	− 1	− 1	− 1	0	− 0,5	0	
L_B	dB		73	72	73,5	73	72,5	73,5	73	72,5	73,5	74	
		a	0	+ 0,5	0	0	+ 1,5	+ 0,5	− 0,5	0	− 0,5	0	
		b	+ 0,5	+ 0,5	0	+ 0,5	+ 1,5	+ 0,5	− 1	0	− 0,5	− 0,5	
NR	Number		65	68	68	71	72	72	72	67,5	70	70	
		a	− 0,5	− 3	− 1	− 5,5	− 3,5	− 2	− 1	− 3	+ 1	− 1	
		b	+ 2	− 2	− 0,5	− 3	− 2	− 2	− 1,5	− 2	+ 1	− 0,5	

a. Deviations from Port's values.
b. Deviations from Lübcke's values.

3.3. Comparison Between the Subjectively and Objectively Obtained Values

3.3.1. Comparison of the Subjective Sound Intensities with the Calculated Values

For the results according to Zwicker, all calculated values lie in an absolute scatter pattern of 5 dB about the subjectively measured values. The average errors of the calculated process for ten noises are outlined in Table IV.

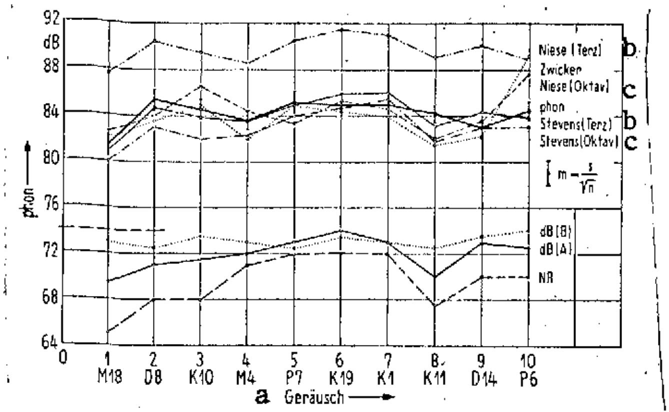

Fig. 10. Graph of the results of all sound intensity evaluation procedures.

Key: a. Third-octave
b. Octave
c. Noise

A higher expenditure, the evaluation of third-octave level diagrams, leads to insignificant deviations from the calculated results, but not always to a greater accuracy than the octave level evaluations. The calculation of the sound intensity from third-octave levels according to Stevens brings the best agreement with the subjective values.

TABLE IV. AVERAGE ERROR IN VARIOUS CALCULATION PROCEDURES.

Procedure	Filter Bandwidth	ΔL dB	s dB
Zwicker	Third	+5.7	±0.0
Stevens	Octave	−1.2	±1
Stevens	Third	0	±0,9
Niese	Octave	+0.2	±1.6
Niese	Third	−0.8	±1.1

One judges the calculation processes according to the variation distribution of the results about the mean deviation; thus the accuracy of the previously described process of Zwicker and Stevens for third-octave level analysis is just as good. The /184 values according to Zwicker nevertheless lie about 6 dB too high. This tendency was confirmed through the findings of the Technical University, Dresden, where sound intensity investigations were run [7, 8]. (The reasons for this effect are not yet clear.)

The values according to Niese are on the average accurate, and the scatter is relatively large. The impulse evaluation does not work out in practice. The measured change in impulses $\Delta L_1 = L_{A1} - L_A$ was ≤ 1 dB.

3.3.2. Comparison of the Subjectively Obtained Sound Intensity with Measured Noise Levels and NR Values

In the measurement of the frequency-evaluated sound pressure levels for the judgment of a sound, very serious errors were made, as shown in Table V.

TABLE V. COMPARISON OF JUDGMENTS WITH LEVELS L_A AND L_B.

Level Difference	$\overline{\Delta L}$ dB	s dB
$L_N - L_A$	12,2	$\pm 1{,}5$
$L_N - L_B$	11	$\pm 1{,}5$

$L_N = L_{1kHz}$

In conclusion, the octave spectra were checked against the noise rating curves. The NR-figures in current use show the same tendency as the A-evaluated level values, but lie on the average 2 dB lower.

These major differences are understandable. They occur without exception around very broad-band noises. The influence of the bandwidth is not considered in the objective measurement process.

4. Comparison of the Calculated and Subjectively Measured Sound Intensities with the Findings of the Stuttgart Technical College and the Berlin Technical University

The sound intensities calculated from the octave levels at the Technical University of Berlin and at the RFZ, according to Stevens, are uniformly precise ($\overline{\Delta L}$ = 0 dB; s = ±0.3 dB). The values calculated at Stuttgart lie around $\overline{\Delta L}$ = 1.2 dB with s = ±1 dB lower.

The agreement with the sound intensities obtained according to Zwicker is not so good. The results of the Technical University of Berlin lie around $\overline{\Delta L}$ = 0.8 dB with s = ±0.5 dB under the corresponding calculated values. The difference of the values obtained by Port is $\overline{\Delta L}$ = -1.5 dB; s = ±0.8 dB. These deviations lead back essentially to the employment of various evaluative models. For each calculation in which the total level lies at 74 dB, the 70 dB model can be used. The third-octave level measured by Lübcke, Mittag and Port, whose peak values lie at 75 to 78 dB and whose mean values lie around 60 dB,were evaluated with the 90 dB model. This curve does not seem to be precise enough for levels of this order of magnitude. Similar evaluations of such spectra with both models show that differences of about 0.5 dB are to be explained thereby. The figure of the loudness-frequency surface was elucidated in principle with the help of a planimeter. A broader explanation offers a somewhat different development of the spectra. The third-octave levels outlined in [1] are smaller up to 4 dB in the range around 2.5 kHz. Here, the hearing perception curves which form the basis of the Zwicker models reach a maximum. This signifies that a difference in this frequency range can be of particularly great weight.

From the comparison of the subjective measurements, one finds as the difference between our results and those of Port

$$L_{NP} = +2.3 \text{ dB}, \; s = \pm 2.1 \text{ dB}$$

between our results and those of Lübcke and Mittag

$$L_{NL} = +0.3 \text{ dB}, \; s = \pm 1.6 \text{ dB}$$

These differences are important under the assumption that between the levels of a 1000-Hz sound and a frequency ensemble noise around 1000 Hz at the same sound intensity, there is a level difference of 2 dB (see Sec. 2.7).

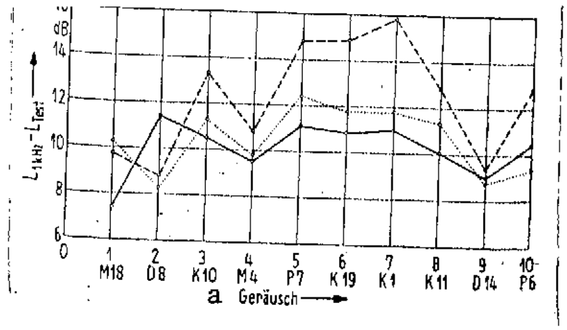

Fig. 11. Comparison of the subjective measurement results at the Stuttgart Technical College, the Technical University of Berlin, and the Berlin RFZ.

– – – Stuttgart
——— RFZ Berlin
----- Techn. Univ. of Berlin

Key: a. Noise

The individual results are displayed in Fig. 11 at any time as the difference between the level of the judged machinery noise and the equally loud 1 kHz sound.

The correspondence of the three results, when one considers the standard deviations for the subjective hearing comparisons discussed in Sec. 2.7.3, is regarded as good or /185
even very good. On the average, however, Port's results are about 2 dB higher, although because of the relatively smaller level of the spectra in the vicinity of best audibility, a negative difference was to be expected.

It is quite difficult to locate possible principles of error sources, when the specified differences lie within the tolerance range of purely accidental errors. The sources for these differences among the subjective findings still were sought in comparing experimental conditions, subjective measurement methods and the measured third-octave levels.

As a first approximation, it was accepted that measurements at very high levels, as indicated by apparent irregularities in the frequency response around 1 kHz, can lead to errors. In such cases, each comparison takes longer to perform, and according to Port and Lübcke, a normal level of 95 dB governs the measurement of the standard sound intensity.

For this presumption, there speaks the fact that the object sound intensities measured at Stuttgart lie on the average 5 dB under the standard sound intensities, while the positive results of a control measurement conducted by Port at L = 40 dB, and the good correspondence of our own subjective sound intensities with the results of measurements of Lübcke and Mittag speak to the contrary.

No basis for the differences in the results can be inferred from the broader objective parameters of the measurement procedures and arrangements used at the various institutions.

The evaluation method still may be a crucial point.

Port has reported the sound intensities based on the value distributions listed by him, not as the arithmetical mean of all individual judgments, but rather has ascertained the median values. Those are the values above which 50% of the measurement values lie, when all values are arranged in numerical order.

In an exact gaussian distribution, the mean values and median values must coincide. With only 12 measurement values, there is the danger that the results for the mean value picture will be relatively greatly falsified by outside values. Relatively large deviations in the results can be avoided through input of the median values. This representation is thus practically a selection of research subjects as well. It must always be considered that deviant results can be confirmed as correct through an enlargement of the observer group.

Both evaluation methods only lead to essentially different values, then, when the so-called outside values and thereby the mean values lie over half of the median values. Here, however, there are positive deviations up to 5 dB in the median values to be explained, and that is not possible with this method. The difference between mean and median values will ordinarily never exceed 0.5 dB.

The subjective influence of different individual sound intensity findings dealt with in Sec. 2.7 remains to be discussed.

Lübcke and Mittag have, as in the publication mentioned earlier [1], selected the observer group painstakingly from a larger number of research subjects.

The diffusion range of the measurements lies between s = ±4 dB and s = ±7 dB.

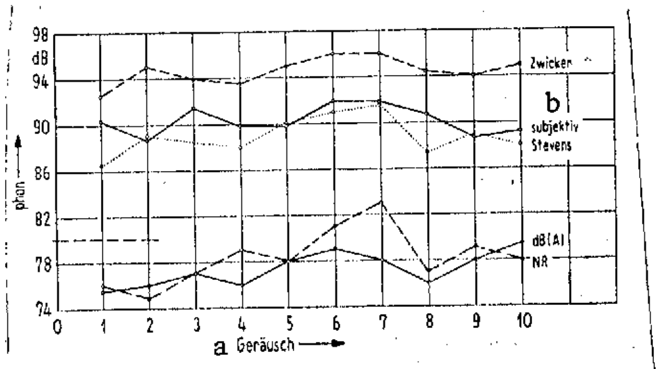

Fig. 12. Graph of the results from the Technical University of Berlin, with $L_{1\ kHz} = L_{normal} - 2$ dB.

Key: a. Noise
b. Subjective

Port restricts himself to the corroboration of 12 research subjects by measurement of absolute hearing thresholds. The diffusion range of his results was appropriately the graph described as "actual variations of the median values" and is thus not directly comparable. For the reported collection specifies the range for the scatter within which 50% of all measurement values lie. The standard deviation encompasses 68% of all measurement values.

5. Conclusions

The exploration of the evaluation procedure at RFZ led to the result that for actual noises of similar type, the calculation of sound intensities from third-octave levels according to the method of Stevens led to the most accurate results. For many practical cases, in which the diffusion range of the subjectively determined loudness values lies from s = ±4 to ±5 dB, the evaluation by octave analysis following Niese or Stevens offers acceptable accuracy. This conclusion is supported by the findings of the Technical University of Berlin, Fig. 12.

P.S.: After the conclusion of our work, we learned of changed /186 Zwicker models (ISO/TC 43 [Secretariat 198] 318/E/August 1963). Through the employment of these new diagram sheets, the difference between subjectively determined sound intensities and the values calculated according to Zwicker, for the ten noises investigated, is lowered by about 1 to 2 dB, on the average by 1.4 dB. The mean error of the procedure was accordingly, for the broad-band noises investigated, L = +4.3 dB. The scatter range of results remains the same.

REFERENCES

1. Lübcke, E., G. Mittag, and E. Port, "Subjective and objective evaluation of machine noise," Acustica 14, 105 (1964).

2. Zwicker, E., G. Flottorp, and S.S. Stevens, "Critical bandwidth in loudness summation," J. Acoust. Soc. Amer. 29, 548 (1957).

3. Niese, H., "Advantages of frequency-modulated normal noise for subjective tests," Lecture delivered at the 3rd Acoustical Conference, 1964, Budapest.

4. Zwicker, E., Method of assessment of loudness from objective analysis, ISO/TC 43 (Helsinki) 35, June 1961.

5. Stevens, S.S., Procedure for calculating loudness, ISO/TC 43 WG 5 (Secretariat-3) 6 February 1961.

6. Niese, H., "A method for determining loudness of random noises," Acustica 15, 117 (1965).

7. Speigel, M., Diploma dissertation, Institute of Electro- and Architectural Acoustics of the Dresden Technical College, 1960.

8. Niese, H., "Loudness of noises and their approximation by measurement and theoretical methods," Hochfrequenztech. u. Elektroak. 72, 3 (1963).

9. Zwicker, E., "Psychological and methodological bases of loudness," Acustica 8, Suppl. 1, 237 (1958).

10. Reichardt, W., "What time constant can reproduce the loudness formation in human hearing?" Lecture delivered at the 3rd Acoustical Conference, 1964, Budapest.